DOCUMENT RESUME

ED 152 134 HE 009 672

AUTHOR Carroll, J. Gregory
TITLE Effects of Training Programs for University Teaching Assistants: A Review of the Literature.
PUB DATE 78
NOTE 57p.; Paper presented at the annual meeting of the American Educational Research Asscciation (Toronto, Ontario, March 1978)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Experiments; *Graduate Students; Higher Education; Institutes (Training Programs); Literature Reviews; Teacher Behavior; *Teacher Education; *Teaching Assistants; *Teaching Skills

ABSTRACT

For this review, approximately 85 studies were identified. The majority were merely descriptions of training programs. There were few studies of teaching assistants' cognitive or affective gains. A number of studies reported significant effects on observed teaching behavior. Whereas quasi-experimental studies generally found significant effects on student variables, true experimental studies generally did not. However, the two types of studies tended to assess different types of training programs. Conclusions, further research questions, and methodological suggestions are offered. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED152134

Effects of Training Programs for University Teaching Assistants: A Review of the Literature

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Gregory Carroll

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

J. Gregory Carroll

University of Pennsylvania

A paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada, March 27-31, 1978.

HE 009 672

Session # 25.23

Abstract

Approximately 85 studies were identified. The majority were merely descriptions of training programs. There were few studies of teaching assistants' cognitive or affective gains. A number of studies reported significant effects on observed teaching behavior.

Whereas quasi-experimental studies generally found significant effects on student variables, true experimental studies generally did not. However, the two types of studies tended to assess different types of training programs. Conclusions, further research questions, and methodological suggestions are offered.

Effects of Training Programs for University Teaching Assistants: A Review of the Literature

As the adequacy of training for college teachers has been criticized more and more widely during the past fifteen or twenty years, countless pre-service orientation programs, in-service workshops, seminars, apprenticeship programs, intern programs, extern programs, and the like have been developed. During roughly the same period, the recruitment of teaching assistants (TAs) has been extended--for both pedagogical and budgetary reasons--to the undergraduate level, and the use of undergraduate TAs may have helped to heighten the perceived need for training as well as to increase the number and diversity of training programs pressed into service.

The recent development of TA training programs seems commendable, but it is surprising to note how little empirical research has been conducted on the effects of such training. Numerous descriptions of innovative training programs have been published, and a number of catalogs for teacher training materials are available (Abbott, Berger, Katz, & Spino, 1973; Houston, 1973a, 1973b; Joyce, Morine, Weil, & Wald, 1971; Prothero, Kosberg, Tribble, Williams, & Luther, 1974). Yet only a handful of empirical studies have included student performance and satisfaction measures to assess the effectiveness of TA training programs. Most published reports either failed to provide evaluative data or merely presented interview or questionnaire data from TAs themselves. A few studies described changes in the attitudes

or teaching behaviors of the TAs but did not include any measure of student variables. Only six studies have been identified that (1) included student variables among the criteria of effectiveness, (2) reported reliable observational data on specific teaching behaviors, and (3) utilized a true experimental design (Campbell & Stanley, 1963).

In view of the wealth of training programs and materials available and the dearth of sound empirical research reported, Rosenshine (Note 1) has concluded that for teacher training in general "We are instruction rich, and data poor" (p. 4). McNeil and Popham (1973) reached much the same conclusion.

> There is a paucity of experimental studies where researchers randomly selected an experimental group from a population of teachers, equipped members of the group with specific performance competencies and then measured the extent to which these teachers (as opposed to controls) both a) performed differently in their classrooms and b) enhanced the cognitive or affective growth of students. [p. 224]

As this review demonstrates, the paucity of such experimental studies is particularly characteristic of research on TA training. The purpose of this review is to assess the empirical evidence for effects of TA training programs on both TAs and their students, to synthesize implications and questions for future investigations, and to make recommendations concerning research methodology.

In all, 48 studies were identified as essentially descriptive accounts of specific training programs. A few of these contained broad summary statements based on questionnaires or interviews, and some reported data on other research questions, but none reported actual data or statistics on effects of TA training. For the most part, these reports simply described the training procedures and, in some cases, mentioned that the participants were generally satisfied or that "complaints about teaching assistants were noticeably reduced from the level encountered prior to the adoption of the program" (Garland, 1969, p. 621). For readers who seek descriptions of selected types of training programs, a list of references may be obtained from the author which identifies the department and institution in which each program has been offered.

The studies that have provided empirical data on the effects of training will be analyzed in the following two sections. The first section is concerned with the research on TA variables: measures of TAs' knowledge, attitudes, and observed teaching behavior. The second section is devoted to student variables: student achievement, attitudes, and ratings of instruction. The final section contains conclusions and recommendations for further research.

Empirical Research on TA Variables

Empirical studies of the effects of training on TA variables vary widely in both methodology and the outcomes measured. Some researchers have measured a variety of training effects by means of carefully controlled experiments while others have simply tabulated interview or

questionnaire feedback from the TAs participating in the training program.

Table 1 lists the seven studies which reported data on TA feedback only. In each case the TAs indicated that they had benefited substantially from the training and that they would highly recommend training for future TAs. Data of this type indicate that the participants were probably well satisfied with the training programs, but these data alone do not yield much valid information on specific changes in the TAs' knowledge, attitudes, or teaching skills which may have occurred as a result of training.

- - - - - - - - - - - - -

Insert Table 1

- - - - - - - - - - - - -

Feedback data are useful for determining whether the training program appeared to meet the perceived needs of its participants but are subject to bias from the possible self-selection of participants and from factors of situational demand (Orne,1962). If offered

as the single criterion of effectiveness, favorable TA feedback does not provide very convincing evidence for the adoption or continuation of training programs.

- - - - - - - - - - - - -

Insert Table 2

- - - - - - - - - - - - -

Table 2 provides a summary of the research on TA knowledge, attitudes, and teaching behavior. The experimental design of each study is categorized as being pre-experimental, quasi-experimental, or true experimental (Campbell & Stanley, 1963), and statistically significant results are summarized.

Effects on TA Knowledge

Table 2 indicates quite clearly that relatively few studies have attempted to measure training effects on TA's cognitive knowledge. Only three studies directly measured TA cognitive outcomes after

training. Kozma, Kulik, Smith, and Weitzer (Note 3) and Weaver and Miller (1975) reported on training programs that included self-instructional materials and quizzes. In both cases the undergraduate TAs were trained to serve as proctors in psychology courses being taught via the Personalized System of Instruction (PSI), as described by Keller (1968). Both studies indicated that the proctors were required to complete each unit of the self-instructional materials and to demonstrate mastery on the self-administered unit quizzes. Neither study compared the quiz performance of the trained proctors to that of a control group or reported any statistical tests of the quiz data.

Anderson (1973) measured several variables by means of a self-report cognitive questionnaire administered at the end of a 15 week training program. The training was designed for TAs serving as classroom teachers, as opposed to PSI proctors, and the scores of the 14 participants were compared to those of a non-equivalent control group. The participants reported that they were significantly more familiar with the use of behavioral objectives and statistical procedures for testing ($p < .05$). The validity of this finding, however, could be questioned because of the possible bias of the participants and possible interactions between selection and other factors, such as history (Campbell & Stanley, 1963).

In sum, very little research has been directed to the effects of TA training programs on TAs' cognitive outcomes, and the data which have been reported are not conclusive.

Effects on TA Attitudes

Only six studies of training effects on TA attitudes have been identified. Rose (1972) reported that a campus-wide program for TA training at the University of California, Los Angeles, resulted in a significant change in the participants' attitudes toward criterion referenced testing ($p < .01$). Another study (Graham, 1971) was an evaluation of an in-service training program for 16 geology TAs at Ohio State University. The author reported a significant difference in the participants' sense of role conflict as measured pre and post by the Role Conflict Assessment Instrument ($p < .01$). Rhyne (1973), however, did not find significant changes in role conflict or in teaching concerns for 12 biology TAs at Georgia State University ($p > .05$). Since these three studies all utilized the one-group pretest-posttest pre-experimental design, the results may not be valid because of history, maturation, testing, instrumentation, and interaction effects (Campbell & Stanley, 1963).

A quasi-experimental study (Anderson, 1973) reported significant differences in attitudes toward the following: teaching skills, behavioral objectives, interaction analysis, testing, and micro-teaching ($p < .05$). The comparison involved 14 participants and 29 non-participants of a one semester course for TAs and faculty at the University of Missouri, Columbia. These results may not be valid because of interaction effects (Campbell & Stanley, 1963).

Yaghlian (1972) also used a quasi-experimental design and found no significant differences in TA interpersonal style or attitudes

toward teaching as a career ($p > .05$). The comparisons were made between participants and non-participants and between participants pre and post. The training program consisted of five workshops for eight chemistry TAs at the University of Michigan.

Finally, Kozma et al., (Note 3) reported no differences in attitudes toward proctoring between trained and untrained proctors in a PSI psychology course at Bowling Green State University. The untrained proctors received an orientation lecture, and the trained proctors completed eight self-instructional units in a training manual. The 24 proctors were randomly assigned to either the training or no-training condition.

Overall, the results on TA attitudes are far from conclusive. Two of the three pre-experimental studies found significant differences in TA attitudes after training. One of the two quasi-experimental studies found significant differences, and the only true experimental study found no differences. One tentative pattern that seems to emerge, however, is that the attitudes which reportedly changed tended to refer to highly specific entities. Changes were reported in attitudes toward criterion referenced testing, behavioral objectives, interaction analysis, testing, microteaching, and the like. Attitudes toward more global concepts such as teaching as a career, teaching concerns, and proctoring, failed to show significant differences. Although the results to date suggest that some TA training programs may be effective for short-term changes in participants' attitudes toward specific teaching methods and

techniques, there are virtually no data on the permanence of such attitude change.

Effects on Teaching Behavior

As can be seen in Table 2, most of the research on TA variables has focused on observation of teaching behavior. In fact, 17 of the 20 studies of TA variables reported data on the TAs' observed teaching behavior.

Four of these studies (Kozma, et al., Note 3; Robin & Cook, 1975; Robin & Heselton, 1977; Weaver & Miller, 1975) involved training programs for PSI proctors. Robin and Cook (1975) utilized a time series, multiple-baseline design to study the effects of individual role-playing sessions on six undergraduate proctors in an education course at the State University of New York, Stony Brook. Undergraduate research assistants used a checklist of 33 items to rate the proctors' performance in giving quizzes. The observational data were not tested for statistical significance, but the authors concluded on the basis of multiple baseline comparisons that for five of the six proctors the training was associated with increases in the overall rate of prescribed proctoring behavior.

Robin and Heselton (1977) also used a multiple baseline design to observe the effects of training on proctoring behavior. In this case, however, 11 undergraduate psychology proctors were randomly assigned to either of two training conditions. One group

of proctors received a self-instructional manual only, while the other proctors participated in group role-playing sessions in addition to receiving the manual. The proctors were observed in both simulated and actual proctoring sessions, and their performance was assessed according to four behavior categories: social behavior, feedback, praise, and prompting. It should be noted, however, that these observations of proctoring behavior were made by one of the authors, and since both authors appear to have conducted the training sessions, the observer probably was aware of the training each proctor had received. Therefore, the results could be biased by experimenter expectancy (Rosenthal, 1968).

The data were not tested for statistical significance, but on the basis of multiple baseline comparisons of simulated proctoring sessions, the authors concluded that for both groups, training was associated with increases in prescribed social behavior, feedback, and praise. Neither group showed an appreciable gain in prompting.

A third multiple baseline study (Weaver & Miller, 1975) involved three undergraduate proctors in behavior analysis at the University of Kansas. The training consisted of completing a self-instructional manual and participating in three individual role-playing sessions. Each proctor was observed by trained raters in 25 simulated proctoring sessions and in one actual proctoring session. The data were not tested for statistical significance, but for all three proctors training was associated

with an increase in the percentage of behavior occurrence from baseline to post-training. Mean baseline percentages from simulated proctoring sessions were 14% for preparation behavior, 52% for prompting and 5% for praise; the corresponding mean percentages after training were 80%, 80%, and 88%. When the proctors were observed working with actual students, the overall mean for the three behaviors was 62%, as compared to an overall mean of 19% before training.

One true experimental study (Kozma, et al., Note 3) investigated the effect of a training manual on proctoring behavior. Observational data of a trained coder and data from student questionnaires on proctoring behavior were analyzed in a study of 24 undergraduate psychology proctors at Bowling Green State University. The proctors were randomly assigned to either an experimental group (i.e., training) or a control group (no training). The results showed no significant differences between groups on either the observational data or the data from student questionnaires (level of significance not specified). The authors noted that although the training manual addressed a range of proctoring behaviors, both the trained and untrained proctors became preoccupied with a single type of behavior -- quiz evaluation.

Among the studies of training programs for TAs serving as classroom teachers (i.e., discussion leaders or laboratory instructors rather than PSI proctors), four utilized pre-experimental research designs to measure effects on teaching behavior: Allen (1976),

Hockett (1972), Rhyne (1973), and Smith (1974). All four researchers utilized the one-group pretest-posttest design so the results of these studies may not be valid because of history, maturation, testing, instrumentation, and interaction effects (Campbell & Stanley, 1963).

Allen (1976) described a training program for biology TAs serving as laboratory instructors at West Virginia University. The author rated the teaching performance of each TA during and after training and found a significant improvement from pre to post ($p < .01$). It should be noted, however, that the rating instrument consisted of high-inference items and that the observer making the ratings was also responsible for conducting the TA training program.

The second pre-experimental study (Hockett, 1972) also examined a training program for laboratory instructors, in this case 16 chemistry TAs and 6 geology TAs at Florida State University. Observational data showed a significantly higher indirect/direct ratio after training than before ($p < .01$). In other words, these data indicate that after training the TAs as a group spent significantly less time lecturing, giving directions, and criticizing students and spent significantly more time accepting feelings, encouraging student input, and using student ideas. There were also significant gains in the number of TA interactions with students and in the frequency of higher level cognitive questions ($p < .05$).

A pre-experimental design was also used to assess the effects of a training program for 12 biology TAs at Georgia State University (Rhyne, 1973). The results of this study also indicated a significant change in the indirect/direct ratio as well as significant changes in the ratio of student talk to teacher talk, the amount of time TAs spent with students in the laboratory, and the number of convergent questions, divergent questions, and total questions asked by the TAs ($p < .05$). There were no significant differences on rhetorical questions or managerial questions ($p > .05$).

The fourth pre-experimental study (Smith, 1974) investigated the effects of a training program for 15 TAs from various departments at the University of Florida. Smith found some decrease in the time spent lecturing and some increase in the time spent asking questions. A significant increase in the time occupied by student responses to questions was reported (level of significance not specified).

The one study of classroom teaching behavior which utilized a quasi-experimental design was that of Anderson (1973). Observational data showed that participants, as compared to non-participants, had significantly more student response, teacher questions, positive reinforcement from the teacher, and silence as well as less lecturing ($p < .05$). There were no significant differences on the use of student ideas, directions to students, negative reinforcement, and confusion ($p > .05$).

A true experimental study (Dalgaard, 1976) was conducted at the University of Illinois, Urbana-Champaign, with 22 TAs in economics, business administration, and geography. The TAs were stratified according to departmental affiliation and randomly assigned to the experimental or control group. Observations were made by trained raters using the Teacher Performance Appraisal Scale. A factor analysis of the data yielded two factors from the 10 items of the rating scale, and these two factors were tested with a two-way analysis of covariance. The results showed significantly higher ratings for the trained TAs than for the controls on both the Aims and Content factor ($p < .0045$) and the Student Involvement factor ($p < .0001$).

A second true experimental study (Arbes & Kitchener, 1974) investigated the effects of training both undergraduate TAs and their students to use the Education Through Student Interaction (ETSI) procedure for class discussions (Kitchener & Hurst, 1974, Note 4). The students were randomly assigned to different types of discussion groups in a psychology course at Colorado State University. Trained observers recorded data on verbal interaction using Bales' (1950) interaction process analysis. The data were not tested for statistical significance, but the authors concluded that groups which were led by a trained TA and/or received group training in ETSI had more students participate in class discussions than groups which did not.

The remaining studies of teaching behavior all used the Flanders (1970) Interaction Analysis Categories (FIAC) to code observational data, and all had true experimental designs: Carroll (1976, 1977), Daniels (1970), Haber (1973), Koffman (1974), Murphy (1972) and Tubb (1974).

The first study (Carroll, 1976, 1977) assessed a training program for psychology TAs at Cornell University. Nineteen TAs were stratified according to measures of verbal aptitude and randomly assigned to either the experimental or control group. The experimental design was the posttest-only control group design (Campbell and Stanley, 1963). Four trained observers utilized FIAC to analyze a videotape of each TA's class. Four other trained raters analyzed the cognitive levels of classroom questions and quiz questions according to the taxonomy of Bloom (1956).

There were no significant differences between groups on the congruity among the cognitive levels of the TAs' classroom questions and quiz questions ($p > .10$), and the FIAC ratings indicated no signficant differences on the ratio of student talk to teacher talk (S/T) ($p > .10$). However, the FIAC data did indicate a significantly higher Indirect/Direct ratio (I/D) for the experimental group than for the control group ($p < .06$). Thus, the trained TAs tended to spend more time than the controls accepting feelings, giving encouragement, using student ideas, and asking questions. They tended to spend less time lecturing, giving directions, and criticizing. Also, data from student questionnaires

showed that the trained TAs made significantly greater use of instructional objectives than the untrained ($p < .07$).

The second study in this group (Daniels, 1970) utilized the pretest-posttest control group design (Campbell & Stanley, 1963). The subjects were eight mathematics TAs at East Carolina University. Five TAs were randomly assigned to the experimental group, three to the control group. Significant differences between the experimental and control groups were obtained in a two-way analysis of variance of the I/D ratio ($p < .05$), but there were no significant differences between groups on the S/T ratio or i/d ratio ($p > .05$).

Haber (1973) reported an experiment involving 12 TAs in business administration at Arizona State University. TAs were randomly assigned to one of two experimental groups or a control group. FIAC data and student ratings on the Purdue Instructor Performance Indicator were obtained for all TAs near the beginning of the semester. The control group received no feedback from the FIAC data or the ratings. The experimental groups were given feedback from the student ratings and FIAC data plus group sessions on FIAC techniques and questioning techniques. The results of data collected after training indicated no significant differences among the groups on the S/T ratio or on four versions of the I/D ratio ($p > .05$). It is not clear whether there were any significant differences among groups on the student ratings, but none is reported.

Koffman (1974) assessed a training program for rhetoric TAs at the University of Massachusetts. The experimental design was once again the pretest-posttest control group design with two experimental groups. One experimental group received instruction plus feedback from student questionnaires and FIAC data. The other experimental group received feedback only, and the control group received neither instruction nor feedback. Koffman analyzed the FIAC matrices of individual TAs in detail but did not test the FIAC data for statistical significance. The author concluded that the FIAC data of all three groups showed observable change from pre to post. The TAs of both experimental groups were judged to have increased their time lecturing, using student ideas, and providing structuring comments, whereas control TAs were judged to have increased their use of student ideas only slightly and to have decreased their lecturing and structuring comments.

Murphy (1972) investigated the differences in teaching behavior attributable to a training seminar for chemistry TAs at Ohio State University. The TAs were assigned at random to either the training or no training condition in a pretest-posttest control group design. FIAC data indicated significant differences in favor of the trained group on the use of praise, encouragement, and questions ($p < .10$). The trained TAs also engaged in significantly less lecturing ($p < .05$). The differences between groups were not significant for student participation, types of questions asked, or proportion of correct student responses ($p > .10$).

Finally, Tubb (1974) assessed the effects of a training program for graduate mathematics TAs at Texas A & M University. Two of the eight TAs were randomly assigned to each of three experimental groups and a control group. The first experimental group was trained only in FIAC techniques during five two-hour sessions. The second experimental group was trained only in Polya's (1945) heuristic questioning strategies (PHQPS) during one introductory meeting and five microteaching sessions. The third experimental group was trained in both FIAC and PHQPS, and the control group received no training.

The I/D ratios were tested by analysis of covariance. For all three experimental groups the I/D ratio increased significantly ($p < .05$), whereas for the control group it did not ($p > .05$).

Overall, the research on TAs' teaching behavior seems to indicate that TA training programs designed to change the participants' teaching in clearly specified directions are generally successful in bringing about an observable change over the period of one or two semesters. The research conducted with PSI proctors to date is by no means definitive, but it does provide evidence that live training sessions and/or self-instructional manuals are associated with the participants' increased use of prescribed techniques of proctoring (Robin & Cook, 1975; Robin & Heselton, 1977; Weaver & Miller, 1975). Moreover, the data suggest that some proctoring skills may be noticeably improved by means of self-instructional materials and practice, while the development of other skills may require live role-playing and

feedback as well (Kozma, et al., Note 3; Robin & Heselton, 1977). The existing data on PSI proctors must be interpreted with caution, however, since the research conducted so far has tended to study very small, highly selective training programs, to utilize quasi-experimental research designs, and to fail to test the results for statistical significance.

The research on TAs serving as classroom teachers is considerably more extensive. Pre-experimental studies (Allen, 1976; Hockett, 1972; Rhyne, 1973; Smith, 1974) and quasi-experimental research (Anderson, 1973) indicated significant behavioral changes in each case. In general, these studies indicated that discussion leaders or laboratory instructors who were trained to engage their students more actively in either class discussions or problem solving did, in fact, show significant changes in comparisons from pre to post. Anderson found significant differences between the trained group and a non-equivalent control group. These results are not conclusive, however, since the research designs did not insure internal validity.

Eight studies were identified that utilized true experimental designs to maximize internal validity. In two of these (Arbes & Kitchener, 1974; Koffman, 1974) the authors failed to test for statistical significance but concluded that observable differences in behavior did result. Of the six studies which did test the observational data statistically only Haber (1973) failed to find any significant differences. Dalgaard (1976) reported significant differences on observations by trained raters using the Teacher

Performance Appraisal Scale. Data obtained from trained raters yielded significant differences on various FIAC measures for Carroll (1976, 1977), Daniels (1970), Murphy (1972), and Tubb (1974).

In sum, the experimental evidence clearly supports the notion that training programs aimed at specific teaching skills can indeed promote short-term changes in the observable teaching behavior of TAs. There are virtually no data, however, on the permanence of these changes. Furthermore, it is not known to what degree these observed behavioral changes are influenced by reactive arrangements (Campbell & Stanley, 1963). Since the TAs know when they are being observed and have at least a rough idea how the researcher wants them to teach, the differences in observed teaching behavior may be artifacts of the observational setting.

It seems logical that teaching behavior ought to be one criterion of success for TA training programs. If the training is intended to improve the participants' teaching skills, then the improvement ought to be directly observable. However, observed teaching behavior should not be considered the only criterion. Rather, it should be viewed, in an experimental sense, as a way of verifying that the treatment (i.e., training) was in fact a treatment. The ultimate goal of changing one's teaching behavior is to enhance student outcomes. Therefore, in measuring the effects of TA training programs, researchers should attempt to link behavioral changes in the TAs' teaching to changes in student outcomes.

Research on Student Variables

The studies listed in Table 3 included measures of student outcomes in evaluating the effects of TA training programs. For the purposes of this review, student variables are categorized into three groups: attitudes, ratings of instruction, and achievement. The term "student attitudes" will generally refer to affective outcomes other than the students' ratings of either their courses or their instructors. For example, attitudinal measures of the students' interest in the subject matter, their preference for a given instructional method, or their feelings about a controversial issue would be considered student attitudes. These attitudes represent affective outcomes from instruction. Students' satisfaction with a particular course or their evaluation of an individual instructor, however, would be considered student ratings of instruction. These attitudes represent affective measures about instruction. Student achievement refers to any test or performance measure of the students' cognitive growth. Also noted in Table 3 are the results of observational measures of teaching behavior, which were discussed previously, in order to associate changes in teaching behavior with changes in student outcomes.

- - - - - - - - - - - -

Insert Table 3

- - - - - - - - - - - -

Inspection of the three columns in Table 3 under the heading "Significant Differences" reveals that student ratings have been employed most frequently as criteria of effectiveness. Of the 21 studies listed, 18 analyzed data on student ratings, 10 reported student achievement data, and 4 included student attitudes, as

defined here. Since many of the studies reported data on at least two categories of student variables, each category will not be reviewed separately here. Studies of PSI proctors will be discussed first. Then the studies of classroom TAs will be grouped according to experimental design.

Research on PSI Proctors

Four of the studies listed in Table 3 were concerned with training PSI proctors: Fernald (1973), Kozma et al. (Note 3), Robin and Heselton (1977), and Saba (Note 6). One of these (Fernald, 1973) was primarily a description of the selection and training procedures for undergraduate psychology proctors at the University of New Hampshire. Also included was a brief summary of several items from student ratings of the proctors. These data, which were obtained with a one-shot case study pre-experimental design (Campbell & Stanley, 1963), indicated that the students rated the proctors highly on overall performance and on their ability to teach complex concepts. It is not known from the data collected, however, to what extent the training program itself influenced student ratings.

As noted in the section on TA variables, Robin and Heselton (1977) utilized a multiple baseline design to evaluate two types of training for undergraduate psychology proctors. One group of randomly assigned proctors received only a self-instructional proctoring manual. The other group received five training sessions in addition to the manual. Student achievement data were obtained for both groups by means of a review test consisting of 12 short

essay questions. For these data, a true experimental comparison was made between the two groups, and no significant differences were found (significance level not specified).

Another study discussed in the section on TA variables (Kozma et al., Note 3) utilized a true experimental design to measure student variables for a proctor training program consisting of a self-instructional manual only. Student achievement was measured by the final examination of the course. Student ratings and student attitudes toward PSI were measured pre and post by means of a questionnaire. No significant differences were found in student achievement, ratings, or attitudes (significance level not specified).

However, Saba (Note 6) incorporated a true experimental design in a study of eight undergraduate PSI proctors at North Georgia College and reported significant differences in student ratings and student achievement. The training consisted of five two-hour sessions in human relations skills -- facilitation, empathy, and interaction analysis. The trainees watched demonstrations, participated in role-playing exercises, and evaluated written case studies. Students in an educational psychology course were randomly assigned to each of the eight proctors. Four proctors were randomly assigned to the training group, four to the control group. Unit test results indicated that the experimental group students needed fewer tests to achieve mastery. The differences between groups were significant for one unit ($p < .01$) but not for

the other seven ($p > .05$). On three units the students of the trained proctors missed significantly fewer items per test, but on the other five units the differences were not significant ($p > .05$). Mean student ratings of the proctors' behavioral characteristics and personal effectiveness were significantly higher for the trained proctors than the controls ($p < .01$), but ratings of the course's educational value showed no significant differences between groups ($p > .05$).

Correlational and Pre-experimental Research

Among the studies of TAs serving as classroom teachers, only Rucker (1963) utilized a correlational design. Rucker analyzed the training programs of 10 departments at the University of Michigan. The training programs were compared according to the degree of structure they were judged to have. The degree of structure was measured by (1) the number of topics discussed in TA orientation programs and (2) the regularity of TA training meetings. These variables were ascertained by means of questionnaires and interviews with TAs and by interviews with TA supervisors. The results showed a slight trend, which was not statistically significant, for the TAs of less structured programs to receive higher mean student ratings on the Purdue Instructor Performance Indicator ($p > .15$).

Four studies of classroom TAs are listed in Table 3 as having utilized strictly pre-experimental research designs: Alexander and Davis (Note 5), Buckenmyer (1972), Prentice, Metcalf, Sharp, and Hard (1976), and Wortman and Hillis (1976). In all four cases

student ratings were the only student variable measured, no tests of the statistical significance of these ratings were reported, and the one-shot case study design was employed. According to Campbell and Stanley (1963), studies with this design "have such a total absence of control as to be of almost no scientific value" (p. 6). The TAs of all four programs received favorable student ratings, but it is not clear to what extent these ratings were attributable to the training per se rather than to other variables. No data were reported for the student ratings of untrained TAs.

Quasi-experimental Research

The following studies from Table 3 are quasi-experimental: Allen (1976), Costin (1968a, 1968b), Hockett (1972), Lewis and Orvis (1973), and Yaghlian (1972). Allen utilized a time series design in an analysis of data selected from student ratings. The training program was conducted for biology TAs at West Virginia University. The training consisted of a three-day orientation program followed by weekly seminars throughout the academic term. Topics included the philosophy of education, student learning, the Bloom (1956) taxonomy, instructional objectives, lecturing, leading discussions, and testing. Classroom visitations and videotape analysis were also incorporated into the training.

On several items from student ratings the author compared the group mean of the trained TAs to the group mean of TAs from five previous semesters (i.e., before the training program began). "These data indicate a consistent increase over the past five semesters in the percentage of students who felt that discussions

with their TAs had been most effective in helping them understand biology" (p. 26). However, the data from student ratings were not tested statistically, and because of the time series design, the possibility of history and instrumentation effects must be kept in mind (Campbell & Stanley, 1963).

The second quasi-experimental study (Lewis & Orvis, 1973) also used a variant of the time series design. This training program was provided for seven TAs in economics at the University of Minnesota. The content of the weekly training seminars included such topics as course design, objectives, testing, student-teacher interaction, discussion techniques, techniques for teaching concepts, and the literature on college teaching and economic education. The same TAs served as the control group during the 1971 fall quarter and as the experimental group during the subsequent winter quarter. Student achievement data and student ratings were obtained at the beginning and end of both quarters.

Student achievement was measure by the Test of Understanding in College Economics, Part 1, Forms A and B. The Purdue Rating Scale for College Instructors was the student ratings instrument. The results indicated significantly higher student achievement on the standardized test for both post-test score ($p < .05$) and gain score ($p < .10$) as well as significantly higher student ratings of instruction after TA training ($p < .01$). The validity of the results, however, could have been influenced by the effect of history.

Yaghlian (1972) employed both the one-group pretest-posttest pre-experimental design and a quasi-experimental design. In this case the training consisted of a workshop on each of five topics: communications, classroom climate, influence styles, cross-ethnic relations, and the use of evaluative feedback. Yaghlian analyzed 10 items from a student evaluation questionnaire administered before and after training for eight chemistry TAs at the University of Michigan. There were significant differences in favor of the post-training measure on the overall mean of the 10 items ($p < .03$). These results were susceptible to the influence of history, maturation, testing, instrumentation, and interaction effects. Significant differences were also obtained on each of the 10 items in a comparison between participants and non-participants ($p < .02$). Assignment to treatment was on a self-selection basis, however, and the participant group could reasonably be expected to have had greater motivation and interest in teaching.

Costin (1968a, 1968b) used a quasi-experimental design to compare the student ratings of participants and non-participants of a seminar entitled "Principles and Methods of Teaching Psychology" at the University of Illinois, Urbana. The topics of the weekly seminars included course design, objectives, testing, teacher-student interactions, ethics in teaching, issues of teaching versus research, and problems involved in conducting research on teaching. The TAs were not assigned to the treatment at random, but rather were required to participate in the seminar unless they

had a scheduling conflict. Costin used analysis of covariance to compare participants and non-participants on end-of-term ratings adjusted for mid-term ratings. On this comparison significant differences were found in favor of the participants on the Rapport factor ($p < .05$) but not on the other four factors: Skill, Structure, Feedback, and Interaction ($p > .05$). At the end of the following semester, two groups of TAs were again compared by analysis of covariance. Both groups had taught for the two consecutive semesters; but whereas one group had participated in the seminar either during the previous semester or during the previous summer term, the other group had not participated. In this comparison the participants were rated significantly higher on the Feedback factor and the Interaction factor ($p < .05$).

The last quasi-experimental study (Hockett, 1972) examined a training program for 6 geology TAs and 16 chemistry TAs at Florida State University. Topics included in the training seminars were the nature of science, a systems model of teaching and learning, objectives, testing, the use of media, questioning techniques, the evaluation of teaching, and self-confrontation via microteaching and video feedback.

Student ratings were measured by the Hildebrand-Wilson Teacher Evaluation Scale, student attitudes by Schwirian's Science Support Scale, and student cognitive gains by the Understanding Science Test. The students of the 22 trained TAs were compared to those of a control group of 12 TAs from chemistry, geology, and biology. The

TAs were not randomly assigned to groups, however. The results showed significant differences in favor of the trained group for student attitudes toward science and student ratings of the course and of the TAs ($p < .05$). There were no significant differences in the cognitive gains of the students ($p > .05$).

The quasi-experimental studies all tend to support the hypothesis that formal courses in TA training can effectively improve the teaching performance of TAs as reflected by student attitudes, ratings, and achievement. There remain serious doubts about the validity of these findings, however, because of apparent methodological weaknesses. The possibility of history or instrumentation effects has already been noted for the two time series studies. The other quasi-experimental studies share the same basic weakness of any nonequivalent control group design. Because the TAs were not randomly assigned to treatment, there is no assurance of a lack of bias between the experimental and control groups initially. TAs who were assigned to the experimental group by self-selection or by similar means could reasonably be expected to have had greater interest in teaching, for example, in the first place. Even though the TAs and students of the two groups were pretested in some cases and no significant differences were found at the outset, the significant differences on posttests could simply be artifacts attributable to a bias in the composition of the groups. It should also be noted that only two of the five studies reported any observational data to assess whether the training actually affected the observable teaching

behavior of the TAs and that these data were merely pre and post measures of the same group.

Research with True Experimental Designs

Among the true experimental studies listed in Table 3, there are two (Arbes & Kitchener, 1974; Boeding & Vattano, 1976) that investigated the effects of training in the Education Through Student Interaction (ETSI) discussion method (Kitchener & Hurst, 1974, Note 4). The ETSI method is a systematic approach to classroom discussion based on a model developed by Hill (1969).

> The discussion technique has three parts and is outlined as follows: (a) an overview of critical definitions and the main theme; (b) a discussion including content analysis, critique of material, integration, and application; and (c) evaluation of the group. [Arbes & Kitchener, 1974, p. 121].

Boeding and Vattano trained 16 undergraduate psychology TAs at Colorado State University. The training consisted of a self-instructional ETSI manual and study guide plus a four-hour workshop on both the ETSI method and conventional lecture-discussion methods (LDM). The participants were all trained in both the ETSI and LDM methods. Students were randomly assigned to TAs, and TAs were randomly assigned to either of two groups. One TA group was instructed to use the ETSI method for the first four weeks of class and LDM for the next four weeks. The other TA group was instructed

to do just the reverse. However, no observational data were obtained to verify that the TAs actually altered their teaching behavior as instructed.

Student questionnaire data indicated a significantly more positive attitude toward LDM than ETSI ($p < .005$). No significant differences were found in the students' interest in the field of psychology, tendency to share ideas or participate, feelings of comfort with the discussion, or tendency to apply material from assigned readings ($p > .05$). Although students reported spending significantly more time preparing for ETSI classes ($p < .01$), there were no significant differences between ETSI and LDM on measures of student achievement ($p > .05$).

Arbes and Kitchener (1974) also investigated the effects of training undergraduate psychology TAs at Colorado State University. Four TAs were trained in weekly sessions with the course instructor to serve as group facilitators and as resource persons for course content. The study also incorporated training in the ETSI method for entire discussion groups. Discussion groups with or without trained TAs and with or without group training were compared.

In general, the data from student questionnaires demonstrated significantly higher ratings for discussion groups with trained TAs, regardless of group training in ETSI, than for the groups with neither TAs nor group training in ETSI ($p < .05$). Most group comparisons of student achievement measures yielded no significant differences ($p > .05$), but there were significant differences on one

test between the groups with trained TAs plus group training in ETSI and the groups that had neither ($p < .05$). Because these comparisons involve groups with trained TAs and other groups without any TAs -- trained or untrained -- the results here are difficult to interpret as to the effects of TA training per se. However, this study does raise intriguing questions about the value of discussion group training as an alternative or a supplement to training for TAs. The data on student achievement and student ratings strongly suggest the need for further research on group training with discussion methods in general and with the ETSI method in particular.

Five true experimental studies of student outcomes from TA training remain: Carroll (1976, 1977), Dalgaard (1976), Daniels (1970), Koffman(1974), and Tubb (1974). Each of these studies included observational data, and four yielded significant differences in observed teaching behavior.

Koffman assessed a training program for rhetoric TAs at the University of Massachusetts according to its effects on teaching behavior, student ratings, and student achievement. The instruments, respectively,were FIAC, a specially constructed 31 item student rating form, and an essay test graded by subject matter specialists. The experimental design was the pretest-posttest control group design (Campbell & Stanley, 1963) with two experimental groups. One experimental group received instruction plus feedback from the questionnaires and FIAC data. The other experimental group received feedback only, and the control group received neither instruction

nor feedback. The instruction was conducted on an individual basis for each TA in the first experimental group. It consisted of approximately 10 hours of consultation and included microteaching exercises, video feedback and critique, analysis of additional student questionnaires, and discussion of selected readings.

Koffman analyzed the FIAC matrices in detail but did not test the data statistically. He concluded that the three groups differed on several FIAC dimensions, as noted in a previous section of this review. The number of significant differences reported for the student ratings was about what would be expected by chance alone from multiple t-tests, but the pattern of significant differences was somewhat suggestive. The feedback-plus-instruction group improved significantly from pre to post on 3 of the 31 items and on 1 of the 10 factors. The feedback-only group declined significantly on 4 of the 31 items, and the control group improved on 1 ($p < .05$). Neither the feedback-only group nor the control group had significant changes on any of the 10 factors ($p > .05$). Finally, there were no significant differences among the three groups on achievement ($p > .05$).

Daniels (1970) also utilized the pretest-posttest control group design. The subjects were eight mathematics TAs at East Carolina University during the fall and winter quarters of 1969-1970. Five TAs were randomly assigned to the experimental group, three to the control group. The training consisted of four one-hour group meetings on FIAC techniques plus two one-hour conferences with each TA to analyze audio recordings of that TA's teaching. Significant differences between the two groups were obtained in a two-way analysis of variance of several FIAC measures ($p < .05$), but there were no

significant differences between groups on student achievement ($p > .05$).

The research of Tubb (1974) involved eight mathematics graduate TAs at Texas A & M University. Two TAs were randomly assigned to each of three experimental groups and a control group. The first experimental group was trained only in FIAC techniques during five two-hour sessions. The second experimental group was trained only in Polya's (1945) heuristic questioning strategies (PHQPS) during one introductory meeting and five microteaching sessions. The third experimental group was trained in both FIAC and PHQPS, and the control group received no training.

Verbal interaction was measured pre and post by FIAC. A 19 item test covering four of the classes taught near the end of the course was administered immediately after the fourth class in order to assess student achievement. Student ratings were obtained before and after training by means of a 35 item questionnaire.

The Indirect/Direct (I/D) ratios derived from the FIAC data were analyzed by analysis of covariance. For all three experimental groups the I/D ratio increased significantly ($p < .05$), whereas for the control group it did not ($p > .05$).

The results of the problem solving items from the achievement test showed that both the PHQPS students and the PHQPS-plus-FIAC students scored significantly higher than the control students ($p < .02$). On the same problem solving items, there were no significant differences between FIAC students and control students

and no significant differences among the students of the three experimental groups ($p > .10$).

On the recall and recognition items of the achievement test, however, both the FIAC students and the FIAC-plus-PHQPS students scored significantly higher than control students ($p < .01$, .06, respectively), but the PHQPS students did not ($p > .10$). The FIAC students also scored significantly higher than the PHQPS students ($p < .01$). Thus for the recall and recognition items, the FIAC training was effective in improving student achievement whether the FIAC was used alone or with PHQPS. The PHQPS training alone was not conclusively effective. For the problem solving items, however, the PHQPS training was effective whether used alone or with FIAC, and the FIAC was not conclusively effective.

The data from student ratings were factor analyzed, and each of the seven factors was tested for significant differences from pre to post. Significant differences in favor of the FIAC-plus-PHQPS group were obtained on the Classroom Discussion factor and the General Teaching Ability factor ($p < .01$).

Dalgaard (1976) assessed the effects of a training program for 22 novice TAs of economics, geography, and business administration at the University of Illinois, Urbana-Champaign. The TAs were stratified by academic department and randomly assigned to either a training group or a control group. The training consisted of six two-hour seminars and an individual videotape critique session with

a trained TA supervisor. The topics of the training seminars included educational objectives, testing, teacher-centered instructional methods, student-centered instructional methods, audio-visual materials, lesson planning, and TA self-evaluation.

Observational data obtained from experienced supervisors indicated significant differences in favor of the trained group on the Aims and Content factor ($p < .0001$) and the Student Involvement factor ($p < .0045$). However, no differences were found between groups on student ratings of the TAs, as measured by the Illinois Course Evaluation Questionnaire.

A study of 19 psychology TAs at Cornell University (Carroll, 1976, 1977) did report significant differences in the student ratings of trained and untrained TAs. The TAs were stratified according to measures of verbal aptitude and were randomly assigned to either a training group or a control group. The training included five two-hour workshops, two individual conferences with the TA supervisor, and an individual videotape critique session. Topics included in the training program were instructional objectives, testing, the structure of knowledge in psychological research, active listening, classroom climate, nonverbal communication, questioning techniques, and student feedback.

Observational data from trained FIAC raters indicated significant differences between groups on the I/D ratio ($p < .06$). Data from student questionnaires indicated that the trained TAs also made significantly greater use of instructional objectives than did controls ($p < .07$).

Student ratings were measured by items from the Cornell Inventory for Student Appraisal of Teaching and Courses and an informal supplementary questionnaire. The trained group received significantly higher ratings than the control group on student ratings of both TA effectiveness and seminar effectiveness ($p < .10$).

Summary

The research on student outcomes of TA training programs has tended to rely most often on student ratings of instruction. Measures of cognitive gains or student attitudes toward variables other than the course or the instructor have been obtained less often.

Four training programs for PSI proctors were assessed. Trained proctors were reported to receive high ratings from students in one study (Fernald, 1973). The other studies all utilized a true experimental design. Proctor training via a self-instructional manual did not significantly increase student achievement, ratings, or attitudes in one case (Kozma, et al., Note 3). Robin and Heselton (1977) found no significant differences in student achievement when proctors given a proctoring manual were compared to proctors who received both the manual and several group training sessions. However, proctor training sessions in human relations skills did significantly increase student ratings of proctors and improve student performance on some unit tests (Saba, Note 6). Thus, the existing research on proctor training is inconclusive about the relative value of live training sessions versus proctor manuals. Instruction in human relations skills, however, appears to be a promising direction for training PSI proctors.

Among the studies of classroom TAs, one correlational study (Rucker, 1963) reported a non-significant trend for the TAs of less structured training programs to receive higher student ratings. Also reported were four pre-experimental studies of student ratings (Alexander & Davis, Note 5; Buckenmyer, 1972; Prentice, et al., 1976; Wortman & Hillis, 1976). In each case the trained TAs received favorable student ratings, but the effects of training on these ratings were not clear.

The training programs assessed by means of quasi-experimental designs (Allen, 1976; Costin, 1968a, 1968b; Hockett, 1972; Lewis & Orvis, 1973; Yaghlian, 1972) all tended to span a broad range of topics such as interpersonal communications, classroom climate, instructional objectives, testing, lecturing, questioning, student-centered teaching, influence styles, student feedback, and the philosophy of education. The results of the quasi-experimental research generally supported the hypothesis that formal courses in TA training can significantly improve the participants' performance according to measures of student attitudes, ratings, and achievement.

By contrast, the training programs assessed in most true experimental studies tended to concentrate on a narrower range of topics. Although the training programs of Carroll (1976, 1977) and Dalgaard (1976) covered a range of topics comparable to the quasi-experimental studies, the other studies tended to focus on

specific techniques such as ETSI (Arbes & Kitchener, 1974; Boeding & Vattano, 1976), FIAC (Daniels, 1970; Tubb, 1974), microteaching skills and the use of student feedback (Koffman, 1974), and heuristic questioning (Tubb, 1974).

Although student attitudes were significantly more positive toward the lecture-discussion method than ETSI, there were no significant differences in other student attitudes or student achievement. Training mathematics TAs in FIAC techniques did not yield significant differences in student achievement in the study by Daniels but did lead to significant differences in student achievement on recall items in Tubb's experiment. Tubb also found that training mathematics TAs in both FIAC techniques and heuristic questioning strategies yielded significant differences in student achievement on problem solving items and recall items and yielded significant differences in student ratings. A training program for rhetoric TAs on microteaching skills and the use of student feedback produced no significant differences in student achievement and only chance differences in student ratings.

Both Dalgaard and Carroll assessed training programs that were general in scope and comparable to the type of programs investigated by the quasi-experimental research. Dalgaard did not obtain significant differences in student ratings of TAs in economics, geography, and business administration. Carroll did find significant differences in student ratings of psychology TAs.

Conclusions and Implications for Future Research

One conclusion that seems apparent from the research reviewed here is that substantially more effort ought to be devoted to assessing the effects of TA training programs rather than simply describing innovative ways of conducting such programs. Although the literature contains many descriptions of training programs, relatively few empirical studies have investigated the effects of training. Very little attention has been paid to the cognitive or affective gains of TAs. A considerable amount of research on observed teaching behavior does generally indicate, however, that programs designed to promote specific teaching skills will indeed generate changes in observable teaching behavior for one or two semesters.

Quasi-experimental research has suggested that training programs covering a broad range of instructional variables will improve student attitudes, ratings, and achievement. However, studies with true experimental designs have tended to evaluate training programs devoted to more specialized skills, and further research is needed to clarify whether the effects demonstrated by quasi-experimental research are artifactual or real. Thus, researchers should strive to assign TAs randomly to experimental conditions, to measure TA variables as well as student variables,and to gather reliable observational data on the TAs' teaching behavior. Whenever possible it is also advisable, in order to gain statistical precision, to

obtain pre-training measures of dependent variables and to employ analysis of covariance to adjust for initial differences between groups.

Future research could also extend the scope of inquiry in several important directions. First, future research should investigate variables that determine the degree to which TAs actually implement the training. The results of Rhyne (1973) showed no correlation between TA scores on the Rokeach Dogmatism Scale, obtained before TA training, and changes in observed teaching behavior as measured by the Interaction Analysis for Science Teaching. Little research has been done on other personality variables, on background characteristics, or on abilities of TAs. Future studies could measure a broad range of such variables in an attempt to determine the circumstances in which different types of TAs tend to implement the training better than others. Understanding some of the variables that determine or mediate the degree of implementation would be highly useful in selecting TAs and providing them with appropriate training.

Second, future research should attempt to measure the long-term effects of TA training. How does training affect the TAs' teaching beyond the semester or the year in which they are trained? How many go on to careers in college teaching? For those who do, how successful are they, and how satisfied are they in their careers?

One study (Alciatore & Eckert, 1972) investigated some of these questions in a long range follow up of 33 Ph.D. candidates who participated in a teaching intern program from 1954 to 1956 at the University of Minnesota. In this case, however, the participants in the program were selected by a committee on the basis of their promise as future college teachers, so the actual effects of the program cannot be separated from the selection effect. Studies of matched or equivalent groups of participants and nonparticipants would be most useful.

Third, future research should attend to the cost-effectiveness of alternative types of training programs. The major economic costs of a typical program would most likely be the time invested by the TA supervisor and the TAs. Are the outcomes of training worth the time and talent invested? Are existing training programs the most cost-effective alternatives for achieving the desired outcomes?

Finally, research on TA training ought to be carried out in a variety of settings with various types of TAs and students. On the one hand, it is important to test the generalizability of results from previous studies by replicating the same types of training programs. There may be interations between types of program and various institutional or individual difference variables. On the other hand, experiments should also be conducted with new training methods and with emphasis on different teaching skills. We have much to learn about the effects of alternative training methods and about the teaching skills which matter most.

Reference Notes

1. Rosenshine, B. Teacher competency research. Unpublished manuscript, University of Illinois, Urbana, Bureau of Educational Research, 1974.

2. Napell, S. M. Videotapes, self-analysis, and seminars: A TA training program at U.C. Berkeley. Unpublished manuscript, University of California, Berkeley, Graduate Assistants Teaching Programs, 1976.

3. Kozma, R. B., Kulik, J. A., Smith, B. B., & Weitzer, W. H. The effectiveness of trained versus untrained proctors in a PSI course. Paper presented at the Third National Conference on Personalized Instruction in Higher Education, Washington, D.C., May 3-4, 1976.

4. Kitchener, K. G., & Hurst, J. C. Education through student interaction manual. Unpublished manuscript, Colorado State University, 1972.

5. Alexander, L., & Davis, R. H. A system training program for graduate teaching assistants. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, March 1970.

6. Saba, R. G. The effectiveness of human relations training for proctors in a mastery based educational psychology course. Paper presented at meetings of the American Personnel and Guidance Association, New York, N. Y., March 23-26, 1975. (ERIC Document Reproduction Service No. ED 110 893)

References

Abbott, A. R., Berger, A. C., Katz, F. L., & Spino, W. D. Florida catalog of competency-based teacher training materials. Chipley, Florida: Panhandle Area Educational Cooperative, 1973.

Alciatore, R., & Eckert, R. E. A college teaching internship program. Improving College and University Teaching, 1972, 20, 93-96.

Allen, R. D. Effective training for teaching assistants. American Biology Teacher, 1976, 38, 24-27.

Anderson, L. A. Effects of a training program for teaching assistants. (Doctoral dissertation, University of Missouri-Columbia, 1973). Dissertation Abstracts International, 1974, 35, 832A. (University Microfilms No. 74-18, 458)

Arbes, B., & Kitchener, K. G. Faculty consultation: A study in support of education through student interaction. Journal of Counseling Psychology, 1974, 21, 121-126.

Bales, R. G. Interaction process analysis. Cambridge, Mass.: Addison-Wesley, 1950.

Barrus, J. L., Armstrong, T. R., Renfrew, M. M., & Garrard, V. G. How I do it: Preparing teaching assistants. Journal of College Science Teaching, 1974, 3, 350-352.

Bloom, B. S. (Ed.) Taxonomy of educational objectives: Cognitive domain. New York: Longmans, Green and Co., 1956.

Boeding, C. H., & Vattano, F. J. Undergraduates as teaching assistants: A comparison of two discussion methods. Teaching of Psychology, 1976, 3, 55-59.

Braddock, R. (Ed.) Reversing the Peter Principle to help inexperienced graduate assistants teach freshman rhetoric. College English, 1970, 32, 45-49.

Buckenmyer, J. A. Preparing graduate assistants to teach. Improving College and University Teaching, 1972, 20, 142-144.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1963.

Carroll, J. G. Effects of a teaching assistant training program on teaching behavior and student ratings of instruction. (Doctoral dissertation, Cornell University, 1976). Dissertation Abstracts International, 1978, 38. (University Microfilms No. 77-28, 194)

Carroll, J. G. Assessing the effectiveness of a training program for the university teaching assistants. Teaching of Psychology. 1977, 4, 135-138.

Costin, F. A graduate course in the teaching of psychology: Description and evaluation. Journal of Teacher Education, 1968, 19, 425-432. (a)

Costin, F. Seminar in the teaching of psychology. Improving College and University Teaching, 1968, 16, 48-49. (b)

Dalgaard, K. A. Some effects of training on teaching effectiveness of untrained university teaching assistants. (Doctoral dissertation, University of Illinois, Urbana-Champaign, 1976). Dissertation Abstracts International, 1977, 37, 6416A-6417A. (University Microfilms No. 77-8968)

Daniels, J. W. Effects of interaction analysis upon teaching assistants and student achievement in introductory college mathematics (Doctoral dissertation, Indiana University, 1970). Dissertation Abstracts International, 1970, 31, 2768A-2769A. (University Microfilms No. 70-25, 186)

Fernald, P. S. The selection and training of undergraduate teaching assistants for large personalized instruction courses. Teaching of Psychology Newsletter, American Psychological Association, 1973 (March), 3-5.

Flanders, N. A. Analyzing teaching behavior. Reading, Mass.: Addison-Wesley, 1970.

Garland, J. K. Training for teaching assistants -- Trial classes and TV taping. Journal of Chemical Education, 1969, 46, 621.

Gilbert, C. P., & McArthur, J. F. In-service teacher preparation of French graduate assistants: Design and evaluation. The French Review, 1975, 48, 508-521.

Golmon, M. E. A teaching seminar for graduate teaching assistants in zoology. American Biology Teacher, 1975, 37, 231-233.

Graham, M. W. Development of an inservice program for geology teaching assistants to reduce role conflict and to improve teaching skills. (Doctoral dissertation, Ohio State University, 1971). Dissertation Abstracts International, 1972, 32, 3663A. (University Microfilms No. 72-4504)

Haber, F. B. The effect of instruction in questioning strategies on teaching assistants' classroom performance. (Doctoral dissertation, Arizona State University, 1973). Dissertation Abstracts International, 1973, 34, 2458A-2459A. (University Microfilms No. 73-21, 888)

Hill, W. F. Learning through discussion. Beverly Hills, Calif.: Sage, 1969.

Hockett, J. C. An examination of changes in teacher behaviors and learner perceptions associated with a program for teaching assistants at the Florida State University: A pilot study. (Doctoral dissertation, Florida State University, 1972). Dissertation Abstracts International, 1973, 33, 3458A-3459A. (University Microfilms No. 73-00, 215)

Houston, W. R. Resources for performance based education. Albany: New York State Education Department, Division of Teacher Education and Certification, 1973. (a)

Houston, W. R. Resources for performance based education, supplement A. Albany: New York State Education Department, Division of Teacher Education and Certification, 1973. (b)

Joyce, B., Morine, G., Weil, M., & Wald, R. (Eds.) Materials for modules: A classification of competency oriented tools for teacher education. Washington, D.C.: U.S. Office of Education, 1971.

Keller, F. S. "Good-bye, teacher..." Journal of Applied Behavior Analysis, 1968, 1, 79-89.

Kitchener, K. G., & Hurst, J. C. Faculty consultation: An emerging role for the counseling psychologist. Journal of Counseling Psychology, 1974, 21, 127-132.

Koffman, M. L. A study of the effects of instructional analysis and feedback on the classroom behavior and student achievement of university teaching assistants (Doctoral dissertation, University of Massachusetts, 1974). Dissertation Abstracts International, 1974, 35, 302A-303A. (University Microfilms No. 74-15, 026)

Lewis, D. R., & Orvis, C. C. A training system for graduate student instructors of introductory economics at the University of Minnesota. Journal of Economic Education, 1973, 5, 38-46.

McNeil, J. D., & Popham, W. J. The assessment of teacher competence. In R. M. W. Travers (Ed.) Second handbook of research on teaching. Chicago: Rand-McNally and Co., 1973.

Monson, C. H. On teaching teaching assistants. Phi Delta Kappan, 1970, 52, 93-94.

Murphy, M. D. The development and assessment of an experimental teacher training program for beginning graduate assistants in chemistry (Doctoral dissertation, Ohio State University, 1972). Dissertation Abstracts International, 1973, 33, 4223A. (University Microfilms No. 73-2083)

Orne, M. On the social psychology of the psychological experiment. The American Psychologist, 1962, 17, 776.

Polya, G. How to solve it. Princeton: Princeton University Press, 1945.

Prentice, E. D., Metcalf, W. K., Sharp, J. G., & Hard, W. L. Training teachers in the anatomical sciences. Journal of Medical Education, 1976, 51, 1006-1009.

Prothero, J., Kosberg, E., Tribble, A., Williams, T., & Luther, J. June 1974 addendum to the Florida catalog of competency-based teacher training materials. Chipley, Florida: Panhandle Area Educational Cooperative, 1974.

Rhyne, P. J. A training program to improve teaching methods and student-teacher interaction of graduate biology teaching assistants (Doctoral dissertation, Georgia State University, 1973). Dissertation Abstracts International, 1974, 34, 4971A. (University Microfilms No. 74-2893)

Robin, A. L., & Cook, D. A. Training proctors for personalized instruction. In Proceedings of the second national conference on personalized instruction in higher education. Washington, D.C.: Center for Personalized Instruction, 1975.

Robin, A. L., & Heselton, P. Proctor training: The effects of a manual versus direct training. Journal of Personalized Instruction, 1977, 2, 19-24.

Rose, C. An in-service program for teaching assistants. Improving College and University Teaching, 1972, 20, 100-102.

Rosenthal, R. Self-fulfilling prophecy. Psychology Today, 1968, 2 (4), 44-51.

Rucker, M. E. G. An exploratory study of the effectiveness of teaching-fellow training programs in selected departments of the College of Literature, Science, and the Arts at the University of Michigan (Doctoral dissertation, University of Michigan, 1963). Dissertation Abstracts, 1963, 24, 1495. (University Microfilms No. 63-6944)

Smith, A. B. A model program for training teaching assistants. Improving College and University Teaching, 1974, 22, 198-200.

Spears, J., & Zollman, D. Orientation for the new teaching assistant -- A laboratory based program. American Journal of Physics, 1974, 42, 1062-1066.

Tubb, G. W. Heuristic questioning and problem solving strategies in mathematics graduate teaching assistants and their students (Doctoral dissertation, Texas A & M University, 1974). Dissertation Abstracts International, 1975, 36, 235A-236A. (University Microfilms No. 75-15, 077)

Weaver, F. H., & Miller, L. K. The effects of a proctor training package on university students' proctoring behaviors. In J. Johnson (Ed.) Behavior research and technology in higher education. Springfield, Ill.: Thomas, 1975.

Wortman, C. B., & Hillis, J. W. Undergraduate taught "minicourses" in conjunction with an introductory lecture course. Teaching of Psychology, 1976, 3, 69-72.

Yaghlian, N. University teaching: The impact of an in-service program for teaching fellows in chemistry (Doctoral dissertation, University of Michigan, 1972). Dissertation Abstracts International, 1973, 33, 6192A. (University Microfilms No. 73-11, 302)

Table 1

Empirical Studies Reporting Data on TA Feedback Only

Authors	Discipline	University
1. Barrus, Armstrong, Renfrew, & Garrard (1974)	Chemistry	Idaho
2. Braddock (1970)	English	Iowa
3. Gilbert & McArthur (1975)	French	Maryland
4. Golmon (1975)	Zoology	Maryland
5. Monson (1970)	MD[a]	Utah
6. Napell (Note 2)	MD	California, Berkeley
7. Spears & Zollman (1974)	Physics	Kansas State

[a]MD = multi-disciplinary

Table 2

Empirical Studies of Training Effects on TA Variables

	Authors	Experimental Design[a] P	Q	T	Significant Differences on Variables Measured Knowledge	Attitudes	Behavior
1.	Allen (1976)	X					Yes
2.	Anderson (1973)		X		Yes	Yes	Yes/No
3.	Arbes & Kitchener (1974)			X			*[b]
4.	Carroll (1976, 1977)			X			Yes/No
5.	Dalgaard (1976)			X			Yes
6.	Daniels (1970)			X			Yes/No
7.	Graham (1971)	X				Yes	
8.	Haber (1973)			X			No
9.	Hockett (1972)	X					Yes
10.	Koffman (1974)			X			*
11.	Kozma, Kulik, Smith, & Weitzer (Note 3)			X	*	No	No
12.	Murphy (1972)			X			Yes/No
13.	Rhyne (1973)	X				No	Yes/No
14.	Robin & Cook (1975)		X				*
15.	Robin & Heselton (1977)		X				*
16.	Rose (1972)	X				Yes	
17.	Smith (1974)	X					Yes/No
18.	Tubb (1974)			X			Yes
19.	Weaver & Miller (1975)		X		*		*
20.	Yaghlian (1972)		X			No	

[a]Key to experimental designs (Campbell & Stanley, 1963): P = pre-experimental design; Q = quasi-experimental design; T = true experimental design

[b]* = Differences not tested for statistical significance

Table 3

Empirical Studies of Training Effects on Student Variables

	Authors	Experimental Design[a] C	P	Q	T	Obs.[b] Data	Significant Differences[c] Att.	Ratings	Ach.
1.	Alexander & Davis (Note 5)		X					*	
2.	Allen (1976)			X		+		*	
3.	Arbes & Kitchener (1974)				X	*		Yes	Yes/No
4.	Boeding & Vattano (1976)				X		Yes/No		No
5.	Buckenmyer (1972)		X					*	
6.	Carroll (1976, 1977)				X	+		Yes	
7.	Costin (1968a, 1968b)			X				Yes/No	
8.	Dalgaard (1976)				X	+		No	
9.	Daniels (1970)				X	+			No
10.	Fernald (1973)		X					*	
11.	Hockett (1972)			X		+	Yes	Yes	No
12.	Koffman (1974)				X	*		Yes/No	No
13.	Kozma, Kulik, Smith, & Weitzer (Note 3)				X	-	No	No	No
14.	Lewis & Orvis (1973)			X				Yes	Yes
15.	Prentice, Metcalf, Sharp, & Hard (1976)		X					*	
16.	Robin & Heselton (1977)				X	*			No
17.	Rucker (1963)	X						No	
18.	Saba (Note 6)				X			Yes/No	Yes/No
19.	Tubb (1974)				X	+		Yes/No	Yes/No
20.	Wortman & Hillis (1976)		X					*	
21.	Yaghlian (1972)		X	X				Yes	

Table 3 (continued)

[a]Key to experimental designs (Campbell & Stanley, 1963):

C = correlational study

P = pre-experimental design

Q = quasi-experimental design

T = true experimental design

[b]Results of observational data:

+ = significant differences reported

- = data tested for statistical significance; no significant differences found

* = data collected but not tested for statistical significance

[c]Student variables:

Att. = student attitudes (other than attitudes toward course or instructor)

Ratings = student ratings of instruction

Ach = student achievement

* = data collected but not tested for statistical significance